



1

# SEQUENCE LISTING

<110> MCININCH, JAMES

<120> COMPUTATIONAL NUCLEIC ACID CODING AND FEATURE ANALYSIS

<130> 16517.308

<140> 10/775,176

<141> 2004-02-11

<150> 09/698,213

<151> 2000-10-30

<160> 4

<170> PatentIn Ver. 3.2

<210> 1

<211> 2165

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> modified\_base

<222> (2042)

<223> a, c, g, t, unknown or other

<220>

<221> modified\_base

<222> (2061)

<223> a, c, g, t, unknown or other

<400> 1

```
tactcaaaaa tatattccat gcttaattag gccggattcg cggtgacgat gcaccaagag 60
cgggtttttcc gagcattgta ggccgtcctc gccacaccgg tgtgatgggt gggatgggac 120
aaaaggatgc ttatgttgga gacgaggctc aatcaaaacg tggatcttg actctgaagt 180
acccaattga gcatggaatt gttaataatt gggatgacat ggagaagatt tggcatcaca 240
ctttctacaa tgagcttcgt gttgcccctg aagaacatcc ggttctcttg accgaagctc 300
ctctcaatcc gaaagctaac cgtgagaaga tgactcagat catgtttgag acattcaata 360
ctcctgctat gtatgttgcc attcaagctg ttctctcact ctatgccagt ggccgtacta 420
ctggtcagta cattaactaca ttctttttat accgtttggg tgaaataaaa ttcgggtttgg 480
ttcgattcga gtttgctctc attattttta ttttggtggg taggtattgt tttggactcc 540
ggagatgggtg tgagccacac ggtaccaatc tacgaggggt atgcacttcc acacgcaatc 600
ctgcgctcttg atcttgacag tcgtgacctc accgaccacc ttatgaaaat cctgacagag 660
cgtgggttact ctttcaccac aactgctgag cgtgagattg ttagagacat gaaggagaag 720
ctctcttaca ttgccttgga ctttgaacaa gagctcgaga cttccaaaac aagctcatcc 780
ggtgagaaga gcttcgagct gccagacggt caagtgcata ccatcggggc agagcgtttc 840
cgatgccctg aagttctggt tcagccatcg atgatcgga tggaaaatcc gggaattcat 900
gaaactactt acaactcaat catgaaatgt gatgtggata tcaggaagga tctttatgga 960
aacattgtgc ttagtggtgg caccacaatg ttcgatggga ttggtgatag gatgagtaaa 1020
gagatcacag cgttggtctc aagcagtatg aacatcaaag tgggtggctcc accggaaagg 1080
aagtacagtg tctggatcgg tggctctatc ttggcttccc tcagtacttt ccagcaggta 1140
aattacttac tatacttaat acataaagtc tattagtgat ttgatgtata aagtgttaca 1200
```

```

aaaatgtggt ccaaatttgc agatgtggat tgcgaaagcg gagtatgatg aatctggacc 1260
gtcaatcgct cacaggaagt gcttctgata aaaagtcacc aagtaaaaca agagcggtaa 1320
aaattttgat atcagttttt caccctgaag ccagttgcta taattactca caacttctct 1380
atttggtgtc ttttattctt gtccctcggt gtccatttta atctcttttt tgcaacaaag 1440
caacttaaaa aaacagagca gtcattaaca gaatgttatt attatatata tgtatacata 1500
ttagtataca cccattattt cattaaaaaca tttatcatat aaggatagga ttctatacat 1560
cgatatattt attttgttga cactattcag cacatgctta tgtcttatct tgtagtata 1620
tgtaaccaa gacaaataat agatgctaca aattgttttc tttgaagcaa aaatttcaat 1680
cttaaaattg tttttttcca ggttacacaa aaaaaacttg tagtttgtaa attttctata 1740
caattttggg gatctcaaca agaacatgaa cttcaacttc tagtcatatg acgacctgag 1800
tctgcgcggc tgtgaatctc tttgctgcag taaatgttta caagtgggtg gttaaattggt 1860
actgattcaa aagctttaag aaatctacac atttcgtgaa attatttagc agacttgata 1920
ttaaaaaatc aggataaaat gactatccaa agacaaatag gactgtttca catgttcccc 1980
tgattcttgt agtcataaac tcatcagcag ttaacttttc tacctcatac acgctcgcaa 2040
tncgtttgga attatcagct ntaatttttc taattctttg gaaattatta gcagctcgat 2100
caaatggggc atggcttctt cttctatctg caactcatct aaactttcca tgaagaaaca 2160
aagct 2165

```

<210> 2

<211> 412

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Predicted  
synthetic amino acid sequence

<220>

<221> SITE

<222> (1)

<223> There is a predicted stop codon present before the  
codon in corresponding SEQ ID NO: 1 that codes for  
residue 1 of the instant sequence.

<220>

<221> SITE

<222> (6)..(7)

<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 6 to 7 of the instant sequence.

<220>

<221> SITE

<222> (13)..(14)

<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 13 to 14 of the instant sequence.

<220>  
<221> SITE  
<222> (34)..(35)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 34 to 35 of the instant sequence.

<220>  
<221> SITE  
<222> (35)..(36)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 35 to 36 of the instant sequence.

<220>  
<221> SITE  
<222> (70)..(71)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 70 to 71 of the instant sequence.

<220>  
<221> SITE  
<222> (126)..(127)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 126 to 127 of the instant sequence.

<220>  
<221> SITE  
<222> (317)..(318)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 317 to 318 of the instant sequence.

<220>  
<221> SITE  
<222> (330)..(331)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 330 to 331 of the instant sequence.

<220>  
<221> SITE  
<222> (356)..(357)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 356 to 357 of the instant sequence.

<220>

<221> SITE

<222> (412)

<223> There is a predicted stop codon present after the codon in corresponding SEQ ID NO: 1 that codes for residue 412 of the instant sequence.

<400> 2

Arg	Phe	Phe	Arg	Ala	Leu	Ala	Val	Leu	Ala	Thr	Pro	Val	Trp	Leu	Gly	1	5	10	15
Trp	Asp	Lys	Arg	Met	Leu	Met	Leu	Glu	Thr	Arg	Leu	Asn	Gln	Asn	Val	20	25	30	
Val	Ser	Leu	Ser	Thr	Gln	Leu	Ser	Met	Glu	Leu	Leu	Ile	Ile	Gly	Met	35	40	45	
Thr	Trp	Arg	Arg	Phe	Gly	Ile	Thr	Leu	Ser	Thr	Met	Ser	Phe	Val	Leu	50	55	60	
Pro	Leu	Lys	Asn	Ile	Arg	Leu	Thr	Glu	Ala	Pro	Leu	Asn	Pro	Lys	Ala	65	70	75	80
Asn	Arg	Glu	Lys	Met	Thr	Gln	Ile	Met	Phe	Glu	Thr	Phe	Asn	Thr	Pro	85	90	95	
Ala	Met	Tyr	Val	Ala	Ile	Gln	Ala	Val	Leu	Ser	Leu	Tyr	Ala	Ser	Gly	100	105	110	
Arg	Thr	Thr	Gly	Gln	Tyr	Ile	Thr	Thr	Phe	Phe	Leu	Tyr	Arg	Ser	Gly	115	120	125	
Asp	Gly	Val	Ser	His	Thr	Val	Pro	Ile	Tyr	Glu	Gly	Tyr	Ala	Leu	Pro	130	135	140	
His	Ala	Ile	Leu	Arg	Leu	Asp	Leu	Ala	Gly	Arg	Asp	Leu	Thr	Asp	His	145	150	155	160
Leu	Met	Lys	Ile	Leu	Thr	Glu	Arg	Gly	Tyr	Ser	Phe	Thr	Thr	Thr	Ala	165	170	175	
Glu	Arg	Glu	Ile	Val	Arg	Asp	Met	Lys	Glu	Lys	Leu	Ser	Tyr	Ile	Ala	180	185	190	
Leu	Asp	Phe	Glu	Gln	Glu	Leu	Glu	Thr	Ser	Lys	Thr	Ser	Ser	Ser	Val	195	200	205	
Glu	Lys	Ser	Phe	Glu	Leu	Pro	Asp	Gly	Gln	Val	Ile	Thr	Ile	Gly	Ala	210	215	220	
Glu	Arg	Phe	Arg	Cys	Pro	Glu	Val	Leu	Phe	Gln	Pro	Ser	Met	Ile	Gly	225	230	235	240

```

Met Glu Asn Pro Gly Ile His Glu Thr Thr Tyr Asn Ser Ile Met Lys
      245                      250                      255

Cys Asp Val Asp Ile Arg Lys Asp Leu Tyr Gly Asn Ile Val Leu Ser
      260                      265                      270

Gly Gly Thr Thr Met Phe Asp Gly Ile Gly Asp Arg Met Ser Lys Glu
      275                      280                      285

Ile Thr Ala Leu Ala Pro Ser Ser Met Lys Ile Lys Val Val Ala Pro
      290                      295                      300

Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly Gly Ser Ile Val Pro Asn
305                      310                      315                      320

Leu Gln Met Trp Ile Ala Lys Ala Glu Tyr Asn Leu Asp Arg Gln Ser
      325                      330                      335

Ser Thr Gly Ser Ala Ser Asp Gln Lys Ser Pro Ser Lys Thr Arg Ala
      340                      345                      350

Val Lys Ile Leu Asn Ser Ser Ala Val Asn Phe Ser Thr Ser Tyr Thr
      355                      360                      365

Leu Ala Ile Arg Leu Glu Leu Ser Ala Leu Ile Phe Leu Ile Ser Leu
      370                      375                      380

Glu Ile Ile Ser Ser Ser Ile Lys Trp Gly Met Ala Ser Ser Ser Ile
385                      390                      395                      400

Cys Asn Ser Ser Lys Leu Ser Met Lys Lys Gln Ser
      405                      410

```

<210> 3

<211> 411

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Predicted  
synthetic amino acid sequence

<220>

<221> SITE

<222> (1)

<223> There is a predicted stop codon present before the  
codon in corresponding SEQ ID NO: 1 that codes for  
residue 1 of the instant sequence.

<220>  
<221> SITE  
<222> (6)..(7)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 6 to 7 of the instant sequence.

<220>  
<221> SITE  
<222> (13)..(14)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 13 to 14 of the instant sequence.

<220>  
<221> SITE  
<222> (34)..(35)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 34 to 35 of the instant sequence.

<220>  
<221> SITE  
<222> (35)..(36)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 35 to 36 of the instant sequence.

<220>  
<221> SITE  
<222> (70)..(71)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 70 to 71 of the instant sequence.

<220>  
<221> SITE  
<222> (126)..(127)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 126 to 127 of the instant sequence.

<220>  
<221> SITE  
<222> (320)..(321)  
<223> There is a predicted stop codon present between  
the codons in corresponding SEQ ID NO: 1 that code  
for residues 320 to 321 of the instant sequence.

```
<220>
<221> SITE
<222> (329)..(330)
<223> There is a predicted stop codon present between
the codons in corresponding SEQ ID NO: 1 that code
for residues 329 to 330 of the instant sequence.
```

```
<220>
<221> SITE
<222> (355)..(356)
<223> There is a predicted stop codon present between
the codons in corresponding SEQ ID NO: 1 that code
for residues 355 to 356 of the instant sequence.
```

```
<220>
<221> SITE
<222> (411)
<223> There is a predicted stop codon present after the
codon in corresponding SEQ ID NO: 1 that codes for
the residue 411 of the instant sequence.
```

<400> 3															
Arg	Phe	Phe	Arg	Ala	Leu	Ala	Val	Leu	Ala	Thr	Pro	Val	Trp	Leu	Gly
1				5					10					15	
Trp	Asp	Lys	Arg	Met	Leu	Met	Leu	Glu	Thr	Arg	Leu	Asn	Gln	Asn	Val
			20					25					30		
Val	Ser	Leu	Ser	Thr	Gln	Leu	Ser	Met	Glu	Leu	Leu	Ile	Ile	Gly	Met
		35					40					45			
Thr	Trp	Arg	Arg	Phe	Gly	Ile	Thr	Leu	Ser	Thr	Met	Ser	Phe	Val	Leu
	50					55					60				
Pro	Leu	Lys	Asn	Ile	Arg	Leu	Thr	Glu	Ala	Pro	Leu	Asn	Pro	Lys	Ala
65					70					75					80
Asn	Arg	Glu	Lys	Met	Thr	Gln	Ile	Met	Phe	Glu	Thr	Phe	Asn	Thr	Pro
				85					90					95	
Ala	Met	Tyr	Val	Ala	Ile	Gln	Ala	Val	Leu	Ser	Leu	Tyr	Ala	Ser	Gly
			100					105					110		
Arg	Thr	Thr	Gly	Gln	Tyr	Ile	Thr	Thr	Phe	Phe	Leu	Tyr	Arg	Ser	Gly
		115					120					125			
Asp	Gly	Val	Ser	His	Thr	Val	Pro	Ile	Tyr	Glu	Gly	Tyr	Ala	Leu	Pro
	130					135					140				
His	Ala	Ile	Leu	Arg	Leu	Asp	Leu	Ala	Gly	Arg	Asp	Leu	Thr	Asp	His
145					150					155					160
Leu	Met	Lys	Ile	Leu	Thr	Glu	Arg	Gly	Tyr	Ser	Phe	Thr	Thr	Thr	Ala
				165					170					175	

Glu Arg Glu Ile Val Arg Asp Met Lys Glu Lys Leu Ser Tyr Ile Ala  
 180 185 190  
 Leu Asp Phe Glu Gln Glu Leu Glu Thr Ser Lys Thr Ser Ser Ser Val  
 195 200 205  
 Glu Lys Ser Phe Glu Leu Pro Asp Gly Gln Val Ile Thr Ile Gly Ala  
 210 215 220  
 Glu Arg Phe Arg Cys Pro Glu Val Leu Phe Gln Pro Ser Met Ile Gly  
 225 230 235 240  
 Met Glu Asn Pro Gly Ile His Glu Thr Thr Tyr Asn Ser Ile Met Lys  
 245 250 255  
 Cys Asp Val Asp Ile Arg Lys Asp Leu Tyr Gly Asn Ile Val Leu Ser  
 260 265 270  
 Gly Gly Thr Thr Met Phe Asp Gly Ile Gly Asp Arg Met Ser Lys Glu  
 275 280 285  
 Ile Thr Ala Leu Ala Pro Ser Ser Met Lys Ile Lys Val Val Ala Pro  
 290 295 300  
 Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly Gly Ser Ile Leu Ala Ser  
 305 310 315 320  
 Gln Met Trp Ile Ala Lys Ala Glu Tyr Asn Leu Asp Arg Gln Ser Ser  
 325 330 335  
 Thr Gly Ser Ala Ser Asp Gln Lys Ser Pro Ser Lys Thr Arg Ala Val  
 340 345 350  
 Lys Ile Leu Asn Ser Ser Ala Val Asn Phe Ser Thr Ser Tyr Thr Leu  
 355 360 365  
 Ala Ile Arg Leu Glu Leu Ser Ala Leu Ile Phe Leu Ile Ser Leu Glu  
 370 375 380  
 Ile Ile Ser Ser Ser Ile Lys Trp Gly Met Ala Ser Ser Ser Ile Cys  
 385 390 395 400  
 Asn Ser Ser Lys Leu Ser Met Lys Lys Gln Ser  
 405 410

<210> 4

<211> 355

<212> PRT

<213> Arabidopsis thaliana



&lt;400&gt; 4

Gly	Asp	Asp	Ala	Pro	Arg	Ala	Val	Phe	Pro	Ser	Ile	Val	Gly	Arg	Pro
1				5					10					15	
Arg	His	Thr	Gly	Val	Met	Val	Gly	Met	Gly	Gln	Lys	Asp	Ala	Tyr	Val
			20					25					30		
Gly	Asp	Glu	Ala	Gln	Ser	Lys	Arg	Gly	Ile	Leu	Thr	Leu	Lys	Tyr	Pro
		35					40					45			
Ile	Glu	His	Gly	Ile	Val	Asn	Asn	Trp	Asp	Asp	Met	Glu	Lys	Ile	Trp
	50					55					60				
His	His	Thr	Phe	Tyr	Asn	Glu	Leu	Arg	Val	Ala	Pro	Glu	Glu	His	Pro
	65				70					75					80
Val	Leu	Leu	Thr	Glu	Ala	Pro	Leu	Asn	Pro	Lys	Ala	Asn	Arg	Glu	Lys
				85					90					95	
Met	Thr	Gln	Ile	Met	Phe	Glu	Thr	Phe	Asn	Thr	Pro	Ala	Met	Tyr	Val
			100					105					110		
Ala	Ile	Gln	Ala	Val	Leu	Ser	Leu	Ala	Ser	Gly	Arg	Thr	Thr	Gly	Gly
		115					120					125			
Ile	Val	Leu	Asp	Ser	Gly	Asp	Gly	Val	Ser	His	Thr	Val	Pro	Ile	Tyr
	130					135					140				
Glu	Gly	Tyr	Ala	Leu	Pro	His	Ala	Ile	Leu	Arg	Leu	Asp	Leu	Ala	Gly
145					150					155					160
Arg	Asp	Leu	Thr	Asp	His	Leu	Met	Lys	Ile	Leu	Thr	Glu	Arg	Gly	Tyr
				165					170					175	
Ser	Phe	Thr	Thr	Thr	Ala	Glu	Arg	Glu	Ile	Val	Arg	Asp	Met	Lys	Glu
			180					185					190		
Lys	Leu	Ser	Tyr	Ile	Ala	Leu	Asp	Phe	Glu	Gln	Glu	Leu	Glu	Thr	Ser
		195					200					205			
Lys	Thr	Ser	Ser	Ser	Val	Glu	Lys	Ser	Phe	Glu	Leu	Pro	Asp	Gly	Gln
	210					215					220				
Val	Ile	Thr	Ile	Gly	Ala	Glu	Arg	Phe	Arg	Cys	Pro	Glu	Val	Leu	Phe
225					230					235					240
Gln	Pro	Ser	Met	Ile	Gly	Met	Glu	Asn	Pro	Gly	Ile	His	Glu	Thr	Thr
				245					250					255	
Tyr	Asn	Ser	Ile	Met	Lys	Cys	Asp	Val	Asp	Ile	Arg	Lys	Asp	Leu	Tyr
			260					265					270		
Gly	Asn	Ile	Val	Leu	Ser	Gly	Gly	Thr	Thr	Met	Phe	Gly	Gly	Ile	Gly
		275					280					285			

Asp Arg Met Ser Lys Glu Ile Thr Ala Leu Ala Pro Ser Ser Met Lys  
290 295 300

Ile Lys Val Val Ala Pro Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly  
305 310 315 320

Gly Ser Ile Leu Ala Ser Leu Ser Thr Phe Gln Gln Met Gln Met Trp  
325 330 335

Ile Ala Lys Ala Glu Tyr Asp Glu Ser Gly Pro Ser Ile Val His Arg  
340 345 350

Lys Cys Phe  
355